

# TEMPORAL DISCOUNTING PREDICTS STUDENT RESPONSIVENESS TO EXCHANGE DELAYS IN A CLASSROOM TOKEN SYSTEM

DEREK D. REED AND BRIAN K. MARTENS

SYRACUSE UNIVERSITY

Typical assessments of temporal discounting involve presenting choices between hypothetical monetary outcomes. Participants choose between smaller immediate rewards and larger delayed rewards to determine how the passage of time affects the subjective value of reinforcement. Few studies, however, have compared such discounting to actual manipulations of reward delay. The present study examined the predictive validity of a temporal discounting procedure developed for use with children. Forty-six sixth-grade students completed a brief discounting assessment and were then exposed to a classwide intervention that involved both immediate and delayed reinforcement in a multiple baseline design across classrooms. The parameters derived from two hyperbolic models of discounting correlated significantly with actual on-task behavior under conditions of immediate and delayed exchange. Implications of temporal discounting assessments for behavioral assessment and treatment are discussed.

*Key words:* delay discounting, self-control, choice, behavioral economics, assessment

Impulsivity is a defining feature of attention deficit hyperactivity disorder (ADHD) in children (Barkley, 1997, 1998). The presence of impulsive behaviors in childhood has been linked to long-term unemployment, school maladjustment, lack of occupational alternatives, and poor parenting as adults (Bloomquist & Schnell, 2002). Impulsivity (or conversely, deficits in self-control) may also be related to more extreme problem behaviors such as eating disorders, substance abuse, and even suicide (see Wenar & Kerig, 2006).

Basic researchers interested in impulsive and irrational decision making have defined impulsivity as a response profile favoring smaller sooner rewards (SSRs) over larger later rewards (LLRs; Rachlin & Green, 1972), or what is

known as *temporal discounting*.<sup>1</sup> Temporal discounting refers to the phenomenon in which rewards lose their subjective value as the delay to their receipt increases (Ainslie, 1974; Madden & Johnson, 2010). Mazur (1987) developed a procedure for measuring temporal discounting in pigeons by investigating the point at which an SSR was chosen over an LLR within a titrating series of comparison trials. Depending on the subject's choice each trial, the LLR was further delayed when choice favored LLRs or delivered more immediately when choice favored SSRs. This procedure was used to determine the point (i.e., the indifference point or subjective value) at which the subject switched from an LLR to an SSR. Plotting the indifference points (i.e., the subjectively discounted values of the LLR) against their delays revealed a hyperbolic function conforming to the model:

$$V_i = \frac{A_i}{1 + kD}, \quad (1)$$

where  $V$  is the subjective value of a specific

This study was conducted by the first author in partial fulfillment of the requirements for the PhD degree in school psychology at Syracuse University. We thank Florence D. DiGennaro-Reed, Tanya L. Eckert, Martin J. Sliwinski, Lauren Axelrod, and Lauren McClenney for their assistance throughout the duration of this project.

Correspondence concerning this article should be addressed to Derek D. Reed, who is now at the Department of Applied Behavioral Science, Dole Human Development Center, 1000 Sunnyside Ave., Room 4048, Lawrence, Kansas 66045 (e-mail: dreed@ku.edu).

doi: 10.1901/jaba.2011.44-1

<sup>1</sup> The reader will note the use of the term *reward* rather than *reinforcer*. In discussions of hypothetical choices, typically employed in discounting research, consequences are not typically delivered as part of the choice paradigm. Thus, we cannot be certain that the choice produces an actual increase in behavior.

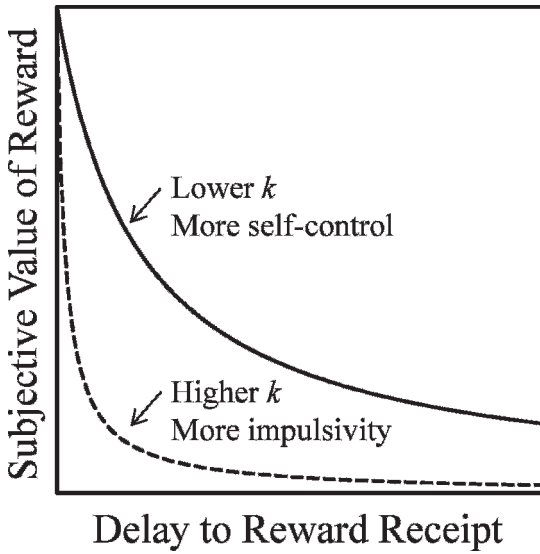


Figure 1. Hypothetical discounting plot depicting both steep discounting (higher  $k$ , dashed curved line; i.e., impulsivity) and less steep discounting (lower  $k$ , solid curved line; i.e., self-control).

reward,  $A$  is the actual amount of the reward,  $D$  is the delay until the reward is obtained, and  $k$  is the derived scaling constant of the hyperbola. This equation proposes a discounting function in which the subjective value of a reward increases as its amount increases, but decreases as a hyperbolic function of the waiting time until it is obtained.

In analyses of discounting, relatively higher  $k$  values translate to relatively higher degrees of impulsive responding, as the hyperbolic curve accelerates downward more rapidly across increasing delays. Figure 1 shows two hypothetical discounting functions, one characteristic of self-control (values discounted less with increases in delay, producing a lower  $k$  value) and one characteristic of impulsivity (values discounted more with increases in delay, producing a higher  $k$  value). According to the model depicted in the equation above, a trade-off exists between the value of a reinforcer and the delay until its receipt. It is important to note that the hyperbolic function depicted by Equation 1 also permits analysis of data patterns that feature preference reversals (i.e., when a

participant chooses the LLR after choosing the SSR at previous delays).

Investigators have typically used hypothetical monetary choice trials to examine the discounting phenomenon with humans. For example, Green, Fry, and Myerson (1994) asked 12 sixth-grade children, 12 college students, and 12 older adults to choose between an SSR and an LLR (each framed as hypothetical monetary rewards) across a series of trials using a titrating procedure adapted from Rachlin, Raineri, and Cross (1991). Specifically, participants were presented with two sets of cards with printed monetary amounts. One set of 30 cards contained the SSR values that varied from 0.1% to 100% of the LLR comparison values. The second set contained the LLR values. Eight delay values were used for each monetary value of the LLR: 1 week, 1 month, 6 months, 1 year, 3 years, 5 years, 10 years, and 25 years. Delays were presented in both progressively ascending and descending orders. Equation 1 was applied to the obtained indifference points, calculated as the average switch point (i.e., when the participant shifted preference from the LLR to the SSR) for each delay between the ascending and descending reward conditions. With an LLR value of \$1,000,  $k$  values for the children, young adults, and older adults were .618, .075, and .002, respectively. The relatively higher  $k$  values suggest more rapidly decreasing slopes (i.e., greater discounting) for the children's discounting plots, suggesting more impulsive responding than the two older groups. Variance accounted for ( $R^2$ ) was .995 for the children, .996 for the young adults, and .995 for the older adults.

Temporal discounting has been used to conceptualize impulsive decisions, such as drug abuse (e.g., Madden, Petry, Badger, & Bickel, 1997), cigarette smoking (e.g., Bickel, Odum, & Madden, 1999), and gambling (e.g., Dixon, Marley, & Jacobs, 2003), as resulting from the lowered subjective value of rewards associated with self-control. This in turn has suggested

procedures for reducing impulsive tendencies by gradually shaping the delay associated with the self-control response. For example, Neef, Bicard, and Endo (2001) examined the relative influence of response effort and reinforcer rate, quality, and immediacy on choice responding of three children with ADHD. Immediacy was the most influential reinforcer dimension in each case, followed by quality and then rate. During subsequent self-control training, delay to receipt of high-quality or high-rate reinforcers was reduced and then systematically increased to the baseline value of 24 hr. By the end of training, all three children exhibited self-control by choosing the reinforcer dimension that competed with immediacy.

Some evidence suggests that results of discounting assessments that use hypothetical rewards, as in Green et al. (1994), are predictive of choices that involve real rewards in laboratory settings (e.g., Johnson & Bickel, 2002; Madden, Begotka, Raiff, & Kastern, 2003; Madden et al., 2004). Moreover, the hypothetical monetary choice trial paradigm has demonstrated adequate reliability as an assessment method for college-aged participants at 3-day (Lagorio & Madden, 2005), 1-week (Simpson & Vuchinich, 2000), 3-month (Ohmura, Takahashi, Kitamura, & Wehr, 2006), and 6-month (Beck & Triplett, 2009) test–retest intervals. Despite promising results in laboratory research, the extent to which discounting assessments are predictive of human behavior in educational settings remains untested. Critchfield and Kollins (2001) have proposed that temporal discounting assessments may be especially advantageous in such settings because they involve behaviors for which consequences are far removed in time or which are indicative of self-control deficits that interfere with contingency learning (e.g., as in children with ADHD). More specific to the current study, because the rewards delivered in educational settings are often delayed (e.g., grades, feedback on performance, token exchanges for primary

reinforcers), a better understanding of the relation between discounting and responsiveness to delayed reinforcers would be beneficial to clinicians and researchers.

The primary goal of this study was to examine the relation between children's responses on hypothetical monetary choice trials and their subsequent responsiveness to both immediate and delayed rewards as part of an independent group-oriented classroom contingency. A secondary goal was to evaluate the consistency of obtained discounting scores across a 1-week interval. In so doing, we sought to determine if a temporal discounting assessment for children demonstrated adequate test–retest reliability and yielded coefficients of stability similar to those found with adult populations (e.g., Beck & Triplett, 2009; Lagorio & Madden, 2005; Ohmura et al., 2006; Simpson & Vuchinich, 2000). Finally, we sought to determine the efficacy of an adapted discounting procedure for children that incorporated shorter delays and smaller reward values. This was done to make the hypothetical choices more similar to the kinds of temporal sequences and monetary amounts with which children may have experience.

## METHOD

### *Participants and Setting*

Students from three sixth-grade classrooms (21, 17, and 8 students from Classrooms 1, 2, and 3, respectively) were recruited for participation ( $n = 46$ ) from a rural public elementary school in the northeast United States. Ages of participants ranged from 11 years 5 months to 12 years 7 months ( $M = 12.1$ ,  $SD = 0.3$ ). Screening criteria for inclusion in the study were (a) the absence of any formal disability classification, (b) English proficiency, and (c) the ability to read, each of which was evaluated through teacher interviews. Consent from the students' primary caregivers, as well as verbal assent from the students themselves, was obtained prior to participation.

During the first phase, participants completed a brief temporal discounting assessment (see below) individually in the hallway outside their classrooms. During the second phase, the experimenter implemented a classwide intervention in the participants' math class. For each of the three classrooms, the intervention occurred during the same math class with the same math teacher (i.e., students rotated from their home classroom to this math classroom).

### *Response Measurement*

*Temporal discounting assessment.* The participants were asked to indicate their preference between two hypothetical choices: one reward that was smaller in magnitude but available immediately (SSR) and a larger reward available after some delay (LLR). During each choice trial, the SSR and LLR were presented opposite one another on a magnetic board (see below). Participants were asked to touch the reward they most preferred during each trial. The reward that the participant touched was considered the preferred choice for that trial (there was never a need for the experimenter to prompt a participant to point or to prompt a participant to point to only one value). The experimenter recorded all choices across all trials on data sheets that displayed all possible choices for each of the eight delay values assessed.

*Classwide intervention.* On-task behavior required that the student have his or her body oriented toward work materials (adapted from Martens, Bradley, & Eckert, 1997). Percentage of intervals scored on task served as the dependent variable across all conditions. On-task behavior was recorded using a 5-s time-sampling procedure, cued by a MotivAider vibrating timer device. Observers began by observing the participant located in one corner of the classroom and moving to the peer adjacent to him or her at the end of each 5-s interval. The observation round was finished after the sequence of observations was completed (i.e., observers recorded data once for each participant using the momentary time-sampling

procedure). Only participants seated at his or her assigned desk at the moment of the time sample were observed. It should be noted that participants usually remained seated in their assigned desks during the duration of the observations. During a 20-min period, observers completed approximately 10 rounds of observation (i.e., obtained approximately 10 time samples per participant).

### *Stimulus Preference Assessment*

A group-administered pictorial-choice preference assessment (adapted from Fisher et al., 1992; see Reed & Martens, 2008) identified preferred academic-related items (pens, pencils, stickers, and erasers) for use as rewards. The stimulus picked most often by the group in each classroom was considered highly preferred, and the stimulus picked least often was considered least preferred. The remaining two items were considered moderately preferred. For Classrooms 1 and 2, pens were highly preferred, with pencils and erasers moderately preferred. For Classroom 3, pencils were highly preferred, and pens and erasers were moderately preferred. For all three classrooms, stickers were the least preferred items.

### *Temporal Discounting Assessment*

The experimenter used a magnetic board (26.2 cm by 35.8 cm) that stood upright on an easel to present hypothetical choices to the participants. Participants sat across the table from the experimenter, with the temporal discounting display board in front of them on the table. Reward values and delays until reward were displayed on magnets (2.6 cm by 10.2 cm) that were placed in their respective positions on the display board. Specifically, amounts displayed on the magnets were in accordance with a titrated series of forced-choice amounts. The experimenter read the following directions to the participants:

Today, we are going to play a pretend game about making choices about money. Since we are only pretending, you will not actually get the money that

you choose. But I want you to pretend that the game is real, so please take your time to decide which amount of money you would really want. Our pretend game will be done with this board. The left side [points to left side] of the board will show a certain amount of money you can pretend to have right now [stress “right now”]. On the right side of the board [points to right side] is the amount of money you can wait [stress “wait”] to have instead. I am going to ask you a lot of pretend decisions to make. When I place the money amounts on the board, I will ask you to pick the one you want the most. If you want to choose the money on the left side, simply point to that side. If you decide that you want to wait to have the money on the right side, simply point to the right side. Let’s practice. Would you rather have \$1 now [puts up \$1 magnet on left], or \$1,000 in 10 minutes [puts up \$1,000 and 10 minutes magnets on right]. Great! You chose —. Now you know how to play. Do you have any questions before we start?

The experimenter manipulated the amount of the SSR at different LLR delay values to determine the points at which each participant changed his or her preference from the LLR (always a hypothetical \$100) to the SSR. Eight LLR delay values were assessed: 1 day, 5 days, 1 month, 2 months, 6 months, 9 months, 1.5 years, and 4 years. These delays remained constant for a block of trials. Within each trial block, the value of the SSR varied depending on each choice made by the participant (see below). Eight indifference points (one at each delay value) were used to fit the hyperbolic curve for each participant. All participants began with the 1-day LLR delay and continued through the series of delays in the same order, from shortest to longest.

The experimenter began each block of trials by asking the participant to choose between \$50 available immediately and \$100 available after a delay of  $X$ , where  $X$  represented the value of the LLR delay during that respective block of trials. If a participant chose the \$50 available now, the next trial pitted an even smaller amount of money available now against \$100 available at the given delay. However, when the participant opted for the \$100 available at the given delay, the experimenter asked the participant to choose between a larger immediate monetary

amount (i.e., SSR) and the delayed \$100 on the next trial. This rapid titration procedure continued until the participant demonstrated preference for the LLR after previously switching to the SSR after an initial LLR choice (or vice versa; see Critchfield & Atteberry, 2003, for a visual depiction). After two such preference reversals, a final trial was conducted to determine the indifference point. If a participant demonstrated exclusive preference for either the LLR or SSR across five consecutive titrations, the discounting assessment trials for that delay value were concluded. The subjective value of the \$100 (i.e., the indifference point) then was derived by averaging the SSR value on the previous trial and the value of the SSR during this final choice trial. This adjusting procedure allowed an estimation of the indifference point (i.e., the averaged SSR value converted to a percentage of LLR; also referred to as the subjective value) of the \$100 across the eight delay values in terms of smaller amounts of money available immediately (ranging from 0.1% to 99.9% of the value of the delayed monetary amount, or \$100) at each delay (see Critchfield & Atteberry, 2003).

The assessment occurred twice during the course of the study, with the two administrations of the assessments separated by 1 week. Repeated administration of the assessment allowed an examination of test–retest reliability over a 1-week interval.

### *Classwide Intervention*

One to 2 days following the second administration of the discounting assessment, the experimenter implemented a classwide intervention targeting on-task behavior in each classroom. Only one session was conducted each day, 5 days per week, during the baseline and sooner reward conditions. During the delayed reward condition, sessions were conducted once per day, Monday through Thursday. The experimenter was present on Friday to deliver rewards earned on Thursday, but no observations were conducted, nor could stu-



dents earn rewards on Fridays during the delayed reward condition. This procedural variation was necessary to ensure that a 24-hr exchange delay was in effect throughout this condition. All sessions lasted 20 min.

*Baseline.* Participants were observed during teacher-led instructional time when the students were expected to be seated at their desks. No programmed reinforcement contingencies were in place for on-task behavior.

*Sooner reward.* Procedures were identical to baseline except that participants could earn a preferred item immediately following the observation. Prior to all intervention sessions, participants were reminded of the classroom token system. The experimenter delivered feedback via a token board to all participants coded as on task during the first, third, fourth, sixth, and eighth rounds of observations. Thus, tokens were delivered immediately following approximately half of the time-sample observations throughout the session. We chose to deliver feedback on an intermittent schedule to reduce the likelihood of reactivity from the students' recognition of the time-sampling intervals. The experimenter provided feedback to participants observed to be on task by placing a token next to the student's name on a display board located on the chalkboard in front of the classroom. Immediately after the observation, participants observed to be on task during all five feedback rounds could choose between one highly preferred item and two moderately or least preferred items. Participants observed to be on task during three or four feedback rounds could choose one moderately preferred item, and participants observed to be on task for fewer than three feedback rounds (but more than zero) could choose one least preferred item. No participant was ever observed to be off task for all five observations.

*Delayed reward.* Procedures were identical to those in the sooner reward condition except that tokens were exchanged for back-up reinforcers at the beginning of the next day's session,

approximately 24 hr after the tokens had been earned. We selected a 24-hr exchange delay to avoid adventitious reinforcement (i.e., exchanging Monday's tokens on Wednesday could adventitiously reinforce responding on Tuesday).

*Design.* The effects of the classwide intervention on on-task behavior were evaluated using a multiple baseline design across classrooms. The design began with three baseline sessions in the first classroom followed by five sessions of the sooner reward condition. After stable data were observed via visual inspection of classroom aggregated data, five sessions of the delayed reward condition occurred for Classroom 1 and seven sessions occurred for Classrooms 2 and 3.

#### *Procedural Fidelity and Interobserver Agreement*

During both temporal discounting assessments and the classwide intervention, a second independent observer was present for at least 33% of each assessment and classroom observation sessions to monitor the fidelity of adherence to the research protocol. During the temporal discounting assessment, the fidelity observer recorded whether the researcher presented the discounting choices in the expected sequence based on participant responses. Deviations from the protocol were scored as a disagreement for that particular block of trials (with a constant large-reward delay) in the temporal discounting assessment. Similarly, procedural fidelity for the classwide intervention was assessed by having a second observer record token and back-up reinforcer delivery. Specifically, the second observer was given a list of participants who met the various reinforcement criteria. This observer then recorded the names of students who obtained reinforcers from the experimenter, as well as those students who did not, along with the number and type of rewards they obtained. This list was cross-referenced with the original list of eligible participants to assess fidelity of both token and reinforcer delivery. In all instances, procedural fidelity was calculated by dividing the

number of instances of agreement (i.e., appropriate implementation of the treatment protocol) by the number of agreements plus disagreements (i.e., instances of deviations from the treatment protocol), multiplied by 100%. For both discounting assessments, as well as during each condition of the classwide intervention, fidelity was 100% for all three classrooms.

A second independent observer was present for at least 33% of the sessions in each classroom to collect data for the purpose of calculating interobserver agreement during each temporal discounting assessment and each condition of the classwide intervention. During the temporal discounting assessment, the second observer recorded the participant's choice on every trial using the same data sheets as the experimenter. Similarly, interobserver agreement for the classwide intervention was assessed by having a second observer independently record student behavior on a classwide intervention data sheet. In all instances, agreement was calculated by dividing the number of instances of agreement by the number of agreements plus disagreements, multiplied by 100%. Interobserver agreement was 100% for all three classrooms for both discounting assessments. During baseline of the classwide intervention, mean agreement was 97% for Classroom 1, 99.6% for Classroom 2, and 97% for Classroom 3. During the sooner reward condition for the three classrooms, mean agreement was 99% for Classroom 1, 99% for Classroom 2, and 100% for Classroom 3. Finally, during the delayed reward condition, mean agreement was 99% for Classroom 1, 99.7% for Classroom 2, and 99% for Classroom 3.

### *Data Analysis*

It has been proposed that adding an additional free parameter (superscript  $s$  in the equation below) to the equation described by Mazur (1987) may account for individual differences in organisms' sensitivity to delay

(i.e., organisms' responses controlled more by reward delay than reward amount; Logue, Rodriguez, Peña-Correal, & Mauro, 1987; Rachlin, 1989). Inclusion of the additional free parameter  $s$  results in the following two-parameter hyperboloid discounting model (Myerson & Green, 1995):

$$V_i = \frac{A_i}{(1 + kD_i)^s} \quad (2)$$

By definition, the inclusion of  $s$  improves the goodness of fit of the discounting model to a data set relative to Equation 1. Many contemporary studies of discounting use this two-parameter hyperboloid model (Equation 2) rather than the simple hyperbolic model in Equation 1 or more complex models of discounting (see McKerchar et al., 2009). Because the field has not conclusively demonstrated superiority of one equation over another, we analyzed our data using both hyperbolic (Equation 1) and hyperboloid (Equation 2) discounting equations. With regard to the use of the hyperboloid model, we examined whether the additional free parameter  $s$  in the two-parameter hyperbolic discounting model offered any advantage over assuming a value of 1 for the students in our sample. This finding would provide further evidence of the robustness of the hyperboloid discounting equation in theoretical discussions of quantitative models and further demonstrate that the findings in basic experimental studies may be translated to elementary-aged students.

Discounting parameters (i.e.,  $k$  values and  $R^2$ ) were obtained by fitting the mean of each participant's indifference points from the two assessment sessions to Equations 1 and 2 using the PROC NONLIN function in the SAS statistical software program. This mean score was used to control for any fluctuations that may have occurred in response patterns between the initial test session and the 1-week retest. The mean across temporal discounting assessments was taken as an estimate of the true score and was used for analyses of the psychometric

properties of the temporal discounting assessment.

Myerson, Green, and Warusawitharana (2001) proposed that, in addition to quantitative models of discounting, researchers also report the area under the discounting curve. Area under the curve (AUC) is a nontheoretical approach to evaluating the degree of discounting, measured by plotting indifference points (i.e., the delay value at which smaller rewards become more preferred than larger rewards) on the ordinate with delays to the larger reward on the abscissa. When thus plotted, low AUC values correspond with high  $k$  values, because the hyperbolic function drops rapidly. Typically, AUC estimates are calculated using the trapezoidal method based on the equation,

$$\text{AUC} = \sum (x_2 - x_1) \left[ \frac{(y_1 + y_2)}{2} \right], \quad (3)$$

where  $x_1$  and  $x_2$  are successive delays and  $y_1$  and  $y_2$  are the subjective value of those delays. This method of analysis was used to provide a theoretically neutral quantification of discounting. Due to the skewed distribution (i.e., variance was not a normally distributed bell curve) of the participants' discounting data, nonparametric Spearman's rho ( $r_s$ ) correlations were conducted to determine the coefficient of stability for participants' indifference points across the 1-week test-retest window.

For all participants, the mean of his or her individual percentage of intervals with on-task behavior during the delayed condition was subtracted from the mean of his or her percentage of intervals on task during the sooner condition to yield a difference score. Correlations then were computed between participants' individual discounting parameters (from Equations 1 and 2 as well as AUC) and their difference scores from the classwide intervention to determine the degree to which the results of the temporal discounting assessment predicted the differential effectiveness of delayed or immediate consequences in the

classwide reinforcement system (i.e., a proxy to predictive validity). Specifically, Spearman's rho correlations were used to determine if those individuals with higher difference scores had lower AUC and higher  $k$  parameters.

## RESULTS

### *Temporal Discounting Assessment*

To determine whether increasing delays were associated with discounted reward values in the hypothetical choices presented to participants, a criterion adapted from Dixon, Jacobs, and Sanders (2006) was employed. Under this criterion, a participant was considered to show a discounting effect if the mean of the indifference points from the three shortest delay conditions exceeded the mean of the indifference points from the three longest delay conditions, with no more than one instance of an increase in indifference points across successive delays (i.e., preference reversal). However, considering the exploratory nature of the current study, this criterion was amended to allow up to two instances of increasing indifference points across successive delays in an effort to maximize the number of participants for analysis. In total, only 26 of the 46 (56%) participants met this inclusionary criterion. It is estimated that discounting studies with adults exclude up to 15% of their data due to invalid patterns (i.e., multiple preference reversals; Critchfield & Atteberry, 2003). Therefore, all analyses and data presented are specific to those participants whose data met the inclusionary criteria. For seven of the 26 participants, we could not fit Equation 2 to the data due to extreme variability in choice.

In strict discounting, each successive delay should feature a lower subjective value (i.e., lower indifference point) than the previous (i.e., a strict monotonically decreasing trend). In the present study, only five participants (P6, P12, P13, P16, and P22) featured such a pattern. However, 16 (61%) of the 26 participants demonstrated a more liberal monotonically



Table 1

Derived  $k$  Values for Equations 1 and 2, Proportions of Variance Accounted for by Equations 1 and 2 ( $R^2$ ), Derived  $s$  Values for Equation 2, Mean Area under the Curve (AUC), and Derived Difference Score from the Classwide Intervention for Each Participant

Participant	Equation 1		Equation 2			Mean AUC	Diff score
	$k$	$R^2$	$k$	$s$	$R^2$		
P1 <sup>a</sup>	0	1.00	−0.14	1.29	.16	.97	.11
P2	0	.99				.69	−.05
P3	0	.98				.65	.14
P4	0	1.00				.44	.04
P5	0	.94	−0.16	1.73	.31	.36	.03
P6 <sup>a</sup>	0.01	.99	0.01	0.02	1.00	.30	.01
P7	0.01	.98				.26	.05
P8	0.02	.95	−0.13	1.23	.36	.24	.13
P9 <sup>a</sup>	0.02	.98	−0.13	1.02	.47	.21	.05
P10	0.01	.99	0.01	0	.99	.17	.16
P11	0.01	.98	0.01	0.11	.99	.16	.06
P12	0.03	.99	0.02	0.07	.99	.12	.10
P13	0.01	.98	0.01	0	.98	.12	−.05
P14	0.02	.98	−0.13	0.93	.56	.11	−.01
P15	0.07	.97	0.04	0.16	.98	.09	.18
P16	0.05	.86	0.02	0.37	.96	.08	.18
P17	0.05	1.00	0.05	0	.99	.07	.07
P18 <sup>a</sup>	0.15	.99	0.16	−0.01	.99	.02	.21
P19 <sup>a</sup>	0.67	.90				.01	.24
P20	1.84	1.00	1.54	0.17	1.00	.01	.22
P21 <sup>a</sup>	0.76	.96	0.23	0.48	1.00	.01	.25
P22	0.93	.99	5.73	−1.31	1.00	.01	.37
P23	1.78	.83	0.15	1.01	.97	.01	.06
P24	0.73	.98	1.21	−0.28	.98	.01	.12
P25	0.87	.96				0	.09
P26	0.61	.86				0	.08

<sup>a</sup> Participant included as an exemplar in Figure 3.

decreasing trend in which each successive delay featured a subjective value that was either less than or equal to the subjective value at the previous delay. The remaining 10 participants (38%) were those individuals with either one or two preference reversals. (Each participant's indifference points across all eight delays from Assessments 1 and 2 are available from the first author.)

Table 1 presents the best fitting values taken by the parameters of Equations 1 and 2 (the discounting models typically employed in discounting studies; see Critchfield & Kollins, 2001; Myerson et al., 2001), as well as the mean AUC and the difference score derived from the classwide intervention, for each of the 26 participants whose data were analyzed (recall that participants with more than two instances

of increasing indifference points as delays increased were excluded from these analyses). The median  $k$  value and  $R^2$  for Equation 1 were .03 (i.e., moderate discounting) and .98 (i.e., 98% of the variance was accounted for by the discounting model) for the 26 participants whose data were analyzed, respectively. Similarly, for Equation 2, the median  $k$  value was .02 and the median  $R^2$  was .98. As noted previously, Equation 2 could not be fit to seven participants' data due to multiple preference reversals. For these participants, only Equation 1 and AUC parameters are reported in Table 1.

Through the use of Equation 2 (i.e., the hyperboloid model), fitting each participant's data to the statistical model not only derived  $k$  and  $R^2$  values but also an exponent value ( $s$ ) that is considered to be a sensitivity-to-delay param-

eter. If the exponent is not found to be significantly different than 1 given the standard error of the fitted parameter, Equation 2 reduces to Equation 1. In the current study, the derived exponents in the hyperboloid model (Equation 2) were indeed significantly different than 1, suggesting that the participants were more sensitive to shorter delays than what would be expected in the hyperbolic model (Equation 1),  $t(17) = -3.65$ ,  $p < .01$ . Thus, a hyperboloid model was necessary to best describe the patterns of responding observed in the discounting assessment for the 19 participants with one or fewer preference reversals. However, for the seven participants who demonstrated two preference reversals, the hyperbolic model was necessary due to the inability of the hyperboloid model to fit their data.

AUC was calculated using Equation 3 for each of the 26 participants using the data from both Assessments 1 and 2. As described previously, the AUC discounting metric is a measure of discounting that is independent of any theoretical model (e.g., hyperbolic vs. hyperboloid). Moreover, AUC ensures that a discounting parameter can be calculated and used to make relative comparisons among participants, even if data do not suggest adherence to discounting assumptions. Thus, AUC may be considered an exact measurement of discounting, unlike the estimations used in Equations 1 and 2. For Assessments 1 and 2, AUC ranged from 0 to .99 and 0 to 1, respectively. These data indicate that our sample was comprised of participants with extreme discounting tendencies (AUC near 0), as well as extremely self-controlled response patterns (AUC near 1).

#### *Test–Retest Reliability*

To investigate test–retest reliability for the temporal discounting assessment, a coefficient of stability (i.e., a correlation coefficient of reliability estimated through the stability of indifference points) was computed for each of

the eight delays, in addition to the AUC values for Assessments 1 and 2. Figure 2 displays the scatterplots of the aggregate indifference points from Session 1 and Session 2, along with the Spearman's rho coefficient of stability ( $r_s$ ), for each of the eight delay values. Specifically, the highest coefficients of stability were found for the following delays: 9 months ( $r_s = .90$ ), 6 months ( $r_s = .87$ ), 5 days ( $r_s = .86$ ), and 1.5 years ( $r_s = .82$ ). The lowest degree of reliability was observed at the 1-day delay ( $r_s = .68$ ). All other delays had reliability coefficients of .81. Thus, adequate levels of test–retest reliability (i.e., .80 and above) were obtained for seven of the eight (87.5%) delay values. Moreover, each of the coefficients at the various delays across all 26 participants was significant at  $p < .01$ . Reliability was assessed for the AUC discounting metric to determine the stability of preferences at the eight delay values across the 1-week test–retest interval. The AUC statistic was also found to be significantly reliable across the 1-week test–retest interval,  $r_s(24) = .88$ ,  $p < .01$ .

#### *Classwide Intervention*

Results of the classwide intervention are presented in Figure 3 (left) for six students, consistent with the multiple baseline design across classrooms (Classrooms A, B, and C). These exemplars were selected based on the criterion that the participant demonstrated either high or low AUC scores relative to other students in the classroom. The multiple baseline design figures in Figure 3 (right) depict discounting plots for Testing Times 1 and 2 as well as AUC estimates for each assessment. Exemplar students with lower AUC scores (the second student in each pair) were considered to be more impulsive, and those with higher AUC scores (the first student in each pair) were considered to be less impulsive than his or her peers (see the discounting plots in the right panels of Figure 3).

As seen in Figure 3 (left), relatively higher levels of on-task behavior were observed during

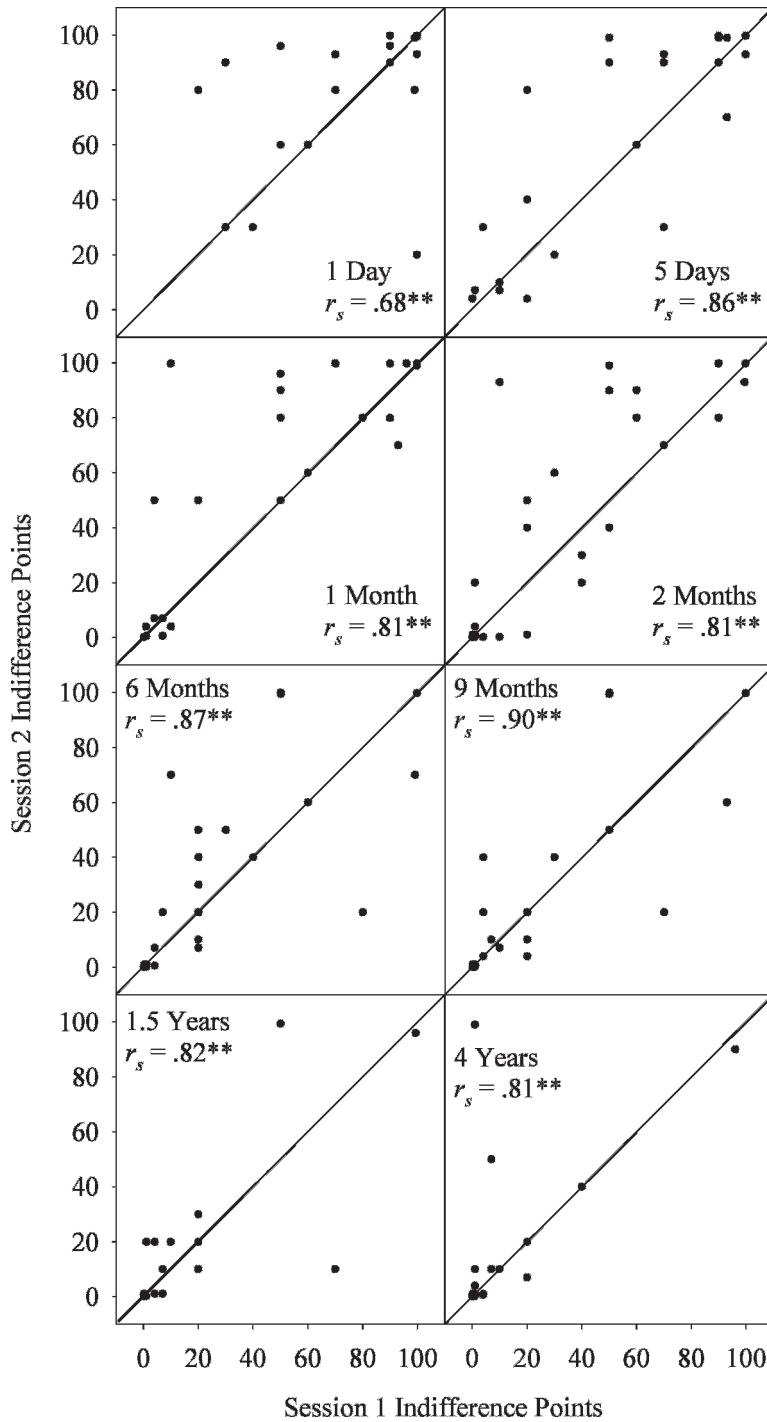


Figure 2. Scatterplots of aggregate indifference points for each of the eight delays, across Sessions 1 and 2. Each panel represents an individual delay value and provides the Spearman's rho correlation for that delay value, along with a diagonal line that indicates the slope of a perfect correlation. Double asterisks indicate that the correlation is significant at the .01 alpha level.

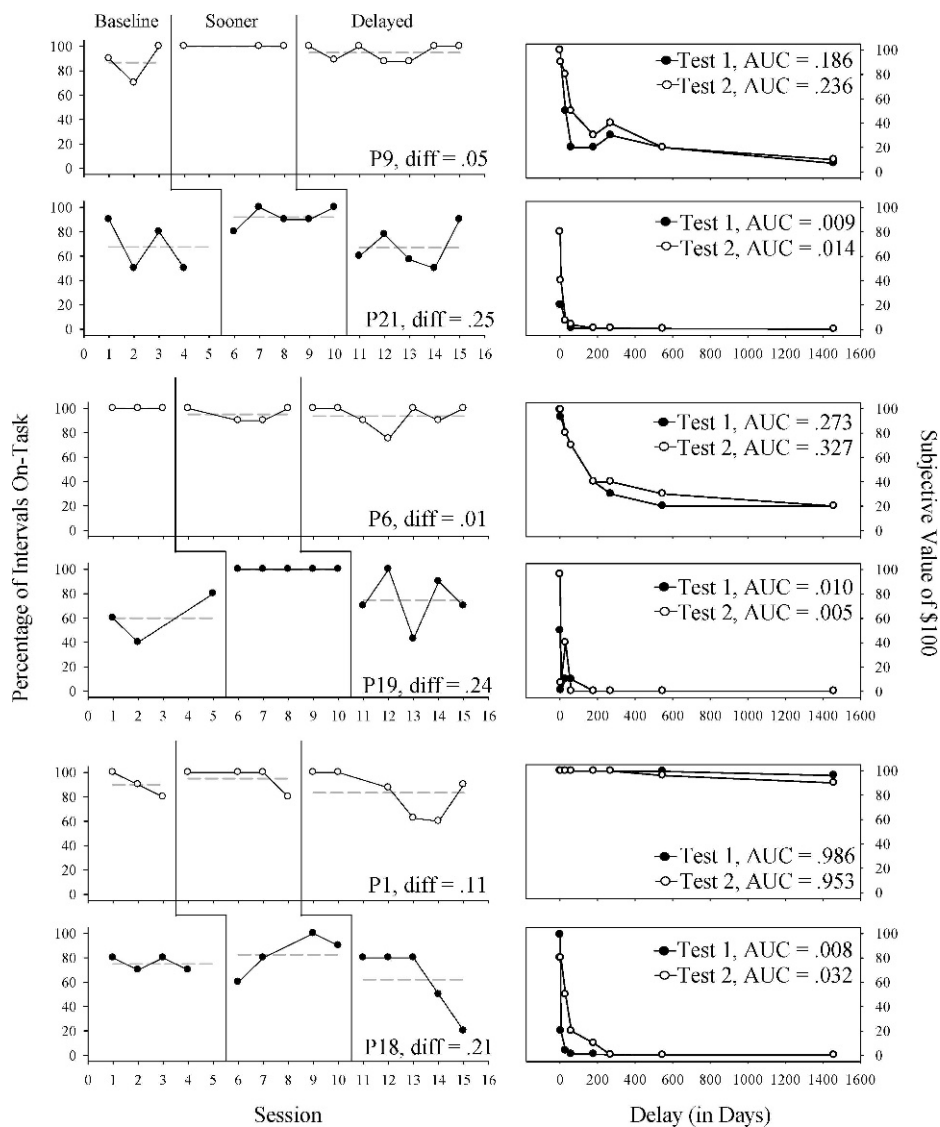


Figure 3. Results of the classwide intervention for on-task behavior in concurrent multiple baseline format for six exemplar students, along with his or her derived difference score between the reinforcement conditions. Filled circles in the left panels represent data for exemplar impulsive students, and open circles represent data for exemplar self-control students. Dashed gray lines represent mean levels of on-task behavior in each condition. Panels to the right of the time-series figures show discounting plots for Testing Times 1 and 2 as well as AUC estimates for each assessment.

baseline for participants who had higher AUC values via the discounting task (Figure 3, right). Increases in on-task behavior during the sooner reward condition relative to baseline were greater for those participants with lower AUC values than for participants with higher AUC values, with P19 having the largest relative

increase (i.e., from approximately 60% on task in baseline to 100% in the sooner reward condition). For two of the three participants with higher AUC values (P1 and P6), only modest improvements were observed in levels of on-task behavior. However, it should be noted that on-task behavior in baseline was relatively

high. For example, P6 had 100% on-task behavior in baseline, so any potential changes during the reinforcement conditions could only be in a negative direction. For all three participants with higher AUC values, transitioning from the sooner reward condition to the delayed reward condition produced only negligible differences in on-task behavior. However, for all three participants with lower AUC, on-task behavior substantially decreased as they transitioned from the sooner to the delayed conditions. Comparing these differences in levels of on-task behavior across reward delay conditions to the AUC scores in Figure 3 (right) illustrates these relations.

Difference scores for all 26 participants ranged from  $-.05$  to  $.37$ , with a median of  $.10$  (see Table 1). In this case, lower scores suggested less difference between immediate and delayed contingencies (or in the case of a negative number, higher percentages in the delayed condition than in the sooner condition), with higher scores indicating more percentage of intervals on task during the sooner condition. Positive difference scores were found for 23 of the 26 cases, suggesting a consistent effect across participants. These difference scores were calculated for use in analyses of the validity of the temporal discounting assessment.

#### *Validity Analyses*

Correlations were computed between pairs of the three discounting parameters (i.e.,  $k$  in Equation 1,  $k$  in Equation 2, and AUC) to determine their degree of correspondence. The correlation between Equation 1  $k$  and AUC was significant and highly negative,  $r_s(24) = -.94$ ,  $p < .01$ , as was the correlation between Equation 2  $k$  and AUC,  $r_s(17) = -.92$ ,  $p < .01$ , using the criterion set forth by Cohen (1988). Thus, higher  $k$  values were indeed associated with lower AUC values. In addition, the correlation between the discounting parameters Equation 1  $k$  and Equation 2  $k$  was significant and highly positive,  $r_s(17) = .90$ ,  $p < .01$ .

Figure 4 shows three scatterplots that depict the relation between participants' discounting parameters and decreases in students' on-task behavior due to reinforcement delay (i.e., the difference score). Each plot features the Spearman's rho linear regression line to describe quantitatively the correlations between these variables. As Figure 4 indicates, the rank-order correlation between each of the three discounting parameters and difference scores was moderate and significant: for Equation 1  $k$  and the difference score,  $r_s(24) = .58$ ,  $p < .01$ ; for Equation 2  $k$  and the difference score,  $r_s(17) = .66$ ,  $p < .01$ ; for AUC and the difference score,  $r_s(24) = -.50$ ,  $p < .01$ , indicating adequate levels of predictive validity. That is, delayed rewards were less effective than immediate rewards for increasing on-task behavior of students with higher discounting scores from the temporal discounting assessment (and lower AUC values), consistent with the data in Figure 3.

## DISCUSSION

Delay of reward is inevitable in most settings. Undoubtedly, no organism experiences immediate contact with rewards at all times. Despite this, many questions remain regarding the analysis of delayed reinforcement (Critchfield & Kollins, 2001). The present study attempted to answer several questions posed by Critchfield and Kollins regarding the use of temporal discounting assessments, thereby replicating and extending previous discounting research. First, does the preference for hypothetical rewards in standard temporal discounting assessments translate to or predict observable behaviors that are reinforced with actual rewards? Similar to previous research (Kirby & Marakovic, 1996; Richards, Zhang, Mitchell, & de Wit, 1999), the present study found adequate levels of predictive validity between a paper-and-pencil discounting assessment for children and their responsiveness to immediate and delayed classroom rewards.



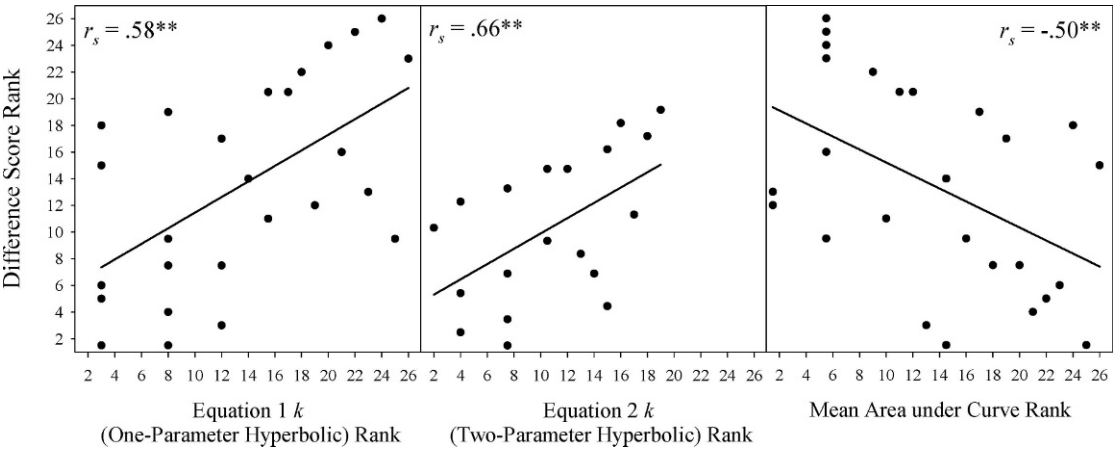


Figure 4. Scatterplots of discounting parameters and derived difference scores from the classwide intervention. Each panel represents an individual discounting parameter and provides the Spearman’s rho correlation for that parameter with the difference score, along with a best fit linear regression line. Double asterisks indicate that the correlation is significant at the .01 alpha level.

Second, are the findings from temporal discounting assessments adequately stable over time? Similar to the previous studies on reliability (Beck & Triplett, 2009; Lagorio & Madden, 2005; Ohmura et al., 2006; Simpson & Vuchinich, 2000), this investigation found adequate test–retest reliability across a 1-week interval for both indifference points and AUC estimates. Thus, the current study suggested that the temporal discounting assessment may be reliable for many children.

Third, are children similar to adults in that they exhibit negative decelerating discounting functions such that the subjective value of the LLR decreases more rapidly at small delays than at longer delays (recall the example in Figure 1)? The present study replicated the findings from Barkley, Edwards, Laneri, Fletcher, and Metevia (2001) and Green et al. (1994) with a majority, albeit a slim one, of responding conforming to general models of discounting. Unlike previous research with children, this study used an adapted assessment (i.e., shorter delays and smaller values) in an effort to make the procedures more applicable to these young participants. Although these efforts were perhaps a step in the right direction, more

adaptations may be warranted given that only 56% of our sample met the discounting criterion adapted from Dixon et al. (2006). In addition, analysis of the additional free parameter  $s$  in Equation 2 found that, similar to adults, the students’ sensitivity to delay was significantly different from 1 (specifically, significantly less than 1), directly replicating the findings of Myerson and Green (1995). Thus, these data suggest that the theoretical models of discounting derived in the basic literature are robust across age groups. Specifically, these data indicate that, similar to adults, children are more sensitive to shorter delays (relative to longer ones) than what is predicted by the simple hyperbolic model of discounting (i.e., Equation 1).

Finally, implications of the present findings may extend beyond the assessment of temporal discounting. The results obtained from the classwide token system offer an interesting insight into the role of token exchange delays in the management of classroom behavior. Past research into exchange delays has suggested that organisms tend to value rewards that are available sooner rather than later (Hackenberg & Vaidya, 2003; Hyten, Madden, & Field,

1994; Jackson & Hackenberg, 1996). To date, these findings have not been replicated in applied settings. The present study showed that students' levels of on-task behavior were sensitive to token exchange delays as part of a classwide token economy. From a clinical standpoint, token-exchange procedures for students with low AUC values (which suggests greater levels of impulsivity) were more likely to be efficacious when exchange delays were relatively brief. The behavior of students with relatively high AUC values improved even when exchange delays were relatively long.

A second implication of these findings is the potential importance of training behavior consistent with self-control when using classroom reinforcement programs. For example, one might establish control over appropriate behavior by immediately delivering large magnitude reinforcers and then gradually increasing the delay across sessions. This approach has been reliably successful for individuals with ADHD (e.g., Binder, Dixon, & Ghezzi, 2000), developmental disabilities (e.g., Vollmer, Borrero, Lalli, & Daniel, 1999), and brain injury (e.g., Dixon, Horner, & Guercio, 2003). A similar fading process might be incorporated into a choice paradigm in which one or more preferred reinforcer dimensions are manipulated to initially bias responding in favor of the delayed alternative, as in Neef et al. (2001). Whether such procedures would ultimately shift indifference points and produce lower  $k$  values during a temporal discounting assessment is a question worthy of future investigation.

A number of procedural limitations compromise the external validity of our findings. Specifically, only 26 of the 46 participants met Dixon et al.'s (2006) criterion for discounting (i.e., consistent decrease in the subjective value of the LLR as delay increases with minimal preference reversals across delays), suggesting a number of future research directions. First, the current methodology employed only hypothetical choices, and our sample of

sixth-grade students may have had little to no prior experience with such choices. Future investigations should examine whether actual reward choice and hypothetical reward choices produce consistent outcomes for this age group. Second, this assessment examined maximum delay values of up to 4 years and maximum reward values of \$100. It seems plausible that reducing these maximum values would have made the hypothetical choices more similar to the actual choices that the participants were likely to have experienced in their day-to-day lives. Thus, more research with children is necessary to isolate the delay and reward values that yield data consistent with discounting assumptions for the majority of participants.

Third, the test-retest reliability of the hypothetical monetary reward task was assessed for only a 1-week interval. The extent to which discounting in children would remain stable over longer time periods given intervening experiences or developmental gains in cognition remains unknown. Fourth, the hypothetical monetary choices in the discounting assessment were qualitatively different from the academic-related rewards used in the classwide intervention, possibly compromising the assumption that one kind of reward could serve as a proxy for another. Fifth, we did not randomize the relative position (left or right) of the SSR, which may possibly confound results if a position bias had developed. (It should be noted that none of the participants was suspected of demonstrating response bias during the assessment [e.g., always choosing the stimulus on the left].) Finally, with the exception of the 1-day delay, the current investigation did not use equivalent delay lengths between the discounting assessment and the exchange delays in the intervention. Future studies that compare subjective values from discounting assessments' delay values to equivalent delay values in actual behavior-change procedures should provide a more direct comparison between hypothetical discounting and actual responsiveness to delayed reinforcement.

In addition to the possible limitations of the temporal discounting assessment, several procedural characteristics of the classwide intervention also limit interpretations of the present findings. First, no reinforcer assessment was conducted to validate the results of the preference assessment due to the large number of students in this study and practical constraints. Thus, we relied on the verbal report of the participants that these rewards were preferred and therefore would function as reinforcers in the classwide token economy. Had we documented that these preferred stimuli were indeed reinforcers, our classroom intervention may have yielded better differentiation across conditions. Second, during baseline of the classwide intervention, levels of on-task behavior for most participants were relatively high, leaving little room to demonstrate effects of the intervention. With such a ceiling effect, we cannot be certain of the potential differences regarding the effectiveness of our delayed and immediate reward conditions.

Third, we did not include formal token training, nor did students in these classrooms have prior experience with token systems. Fourth, all classrooms proceeded through the design in the same order (i.e., baseline, sooner reward, delayed reward). It is possible that order effects may have contributed to the observed changes in behavior during the delayed condition. In addition, the absence of within-subject replication limits the degree to which we could attribute behavior change to the exchange delays alone. This should be addressed in future research. Fifth, the 24-hr delay in the delayed reward condition was arbitrarily chosen in an effort to make the contingencies salient without having additional observations precede the delivery of back-up reinforcers for previous sessions. A longer exchange delay may have produced larger differences between the sooner and delayed reward conditions. Parametric analyses of differing delay values or the use of progressive exchange delays in actual classroom

interventions would better translate the basic research on discounting and exchange delays.

The limitations of the current study notwithstanding, this investigation demonstrated that a hypothetical monetary-choice temporal discounting assessment can yield estimates regarding the degree to which delay of a reward devalues it in a choice task. The study extends previous research in this area by (a) evaluating the psychometric properties of a child-adapted temporal discounting assessment procedure, (b) experimentally manipulating exchange delays to derive predictive validity estimates, and (c) demonstrating that greater degrees of discounting (i.e., relatively more impulsive responding) were correlated with reduced efficacy of delayed back-up reinforcers in a token economy. This investigation suggests that intervention agents may indeed benefit from consideration of the discounting phenomenon in the design and implementation of applied behavior-change procedures.

## REFERENCES

- Ainslie, G. W. (1974). Impulse control in pigeons. *Journal of the Experimental Analysis of Behavior*, 21, 485–489.
- Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive function: Constructing a unified theory of ADHD. *Psychological Bulletin*, 121, 65–94.
- Barkley, R. A. (1998). *Attention-deficit/hyperactivity disorder: A handbook for diagnosis and treatment* (2nd ed.). New York: Guilford.
- Barkley, R. A., Edwards, G., Laneri, M., Fletcher, K., & Metevia, L. (2001). Executive functioning, temporal discounting, and sense of time in adolescents with attention deficit hyperactivity disorder (ADHD) and oppositional defiant disorder (ODD). *Journal of Abnormal Child Psychology*, 29, 541–556.
- Beck, R. C., & Triplett, M. F. (2009). Test-retest reliability of a group-administered paper-pencil measure of delay discounting. *Experimental and Clinical Psychopharmacology*, 17, 345–355.
- Bickel, W. K., Odum, A. L., & Madden, G. L. (1999). Impulsivity and cigarette smoking: Delay discounting in current, never, and ex-smokers. *Psychopharmacology*, 146, 447–454.
- Binder, L. M., Dixon, M. R., & Ghezzi, P. M. (2000). A procedure to teach self-control to children with attention deficit hyperactivity disorder. *Journal of Applied Behavior Analysis*, 33, 233–237.

- Bloomquist, M. L., & Schnell, S. V. (2002). *Helping children with aggression and conduct problems: Best practices for intervention*. New York: Guilford.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Critchfield, T. S., & Atteberry, T. (2003). Temporal discounting predicts individual competitive success in a human analogue of group foraging. *Behavioural Processes*, 64, 315–331.
- Critchfield, T. S., & Kollins, S. H. (2001). Temporal discounting: Basic research and the analysis of socially important behavior. *Journal of Applied Behavior Analysis*, 34, 101–122.
- Dixon, M. R., Horner, M. J., & Guercio, J. (2003). Self-control and the preference for delayed reinforcement: An example in brain injury. *Journal of Applied Behavior Analysis*, 36, 371–374.
- Dixon, M. R., Jacobs, E. A., & Sanders, S. (2006). Contextual control of delay discounting by pathological gamblers. *Journal of Applied Behavior Analysis*, 39, 413–422.
- Dixon, M. R., Marley, J., & Jacobs, E. A. (2003). Delay discounting by pathological gamblers. *Journal of Applied Behavior Analysis*, 36, 449–458.
- Fisher, W., Piazza, C. C., Bowman, L. G., Hagopian, L. P., Owens, J. C., & Slevin, I. (1992). A comparison of two approaches for identifying reinforcers for persons with severe and profound disabilities. *Journal of Applied Behavior Analysis*, 25, 491–498.
- Green, L., Fry, A. F., & Myerson, J. (1994). Discounting of delayed rewards: A life-span comparison. *Psychological Science*, 5, 33–36.
- Hackenberg, T. D., & Vaidya, M. (2003). Determinants of pigeons' choices in token-based self-control procedures. *Journal of the Experimental Analysis of Behavior*, 70, 207–218.
- Hyten, C., Madden, G. J., & Field, D. P. (1994). Exchange delays and impulsive choice in adult humans. *Journal of the Experimental Analysis of Behavior*, 62, 225–233.
- Jackson, K., & Hackenberg, T. D. (1996). Token reinforcement, choice, and self-control in pigeons. *Journal of the Experimental Analysis of Behavior*, 66, 29–49.
- Johnson, M. W., & Bickel, W. K. (2002). Within-subject comparison of real and hypothetical money rewards in delay discounting. *Journal of the Experimental Analysis of Behavior*, 77, 129–146.
- Kirby, K. N., & Marakovic, N. N. (1996). Delay discounting probabilistic rewards: Rates decrease as amounts increase. *Psychonomic Bulletin and Review*, 3, 100–104.
- Lagorio, C. H., & Madden, G. J. (2005). Delay discounting of real and hypothetical rewards: III. Steady-state assessment, forced-choice trials, and all real rewards. *Behavioural Processes*, 69, 173–187.
- Logue, A. W., Rodriguez, M. L., Peña-Correal, T. E., & Mauro, B. C. (1987). Quantification of individual differences in self-control. In M. L. Commons, J. E. Mazur, J. A. Nevin, & H. Rachlin (Eds.), *Quantitative analyses of behavior: Vol. V. The effect of delay and of intervening events on reinforcement value* (pp. 245–265). Hillsdale, NJ: Erlbaum.
- Madden, G. J., Begotka, A. M., Raiff, B. R., & Kastern, L. L. (2003). Delay discounting of real and hypothetical rewards. *Experimental and Clinical Psychopharmacology*, 11, 139–145.
- Madden, G. J., & Johnson, P. S. (2010). A delay discounting primer. In G. J. Madden & W. K. Bickel (Eds.), *Impulsivity: The behavioral and neurological science of discounting* (pp. 1–37). Washington, DC: American Psychological Association.
- Madden, G. J., Petry, N., Badger, G. J., & Bickel, W. K. (1997). Impulsive and self-control choices in opiate dependent patients and non-drug-using control participants: Drug and monetary rewards. *Experimental and Clinical Psychopharmacology*, 5, 256–262.
- Madden, G. J., Raiff, B. R., Lagorio, C. H., Begotka, A. M., Mueller, A. M., Hehli, D. J., et al. (2004). Delay discounting of potentially real and hypothetical rewards: II. Between- and within-subject comparisons. *Experimental and Clinical Psychopharmacology*, 12, 251–261.
- Martens, B. K., Bradley, T. A., & Eckert, T. L. (1997). Effects of reinforcement history and instruction on the persistence of academic engagement. *Journal of Applied Behavior Analysis*, 30, 569–572.
- Mazur, J. E. (1987). An adjusting procedure for studying delayed reinforcement. In M. L. Commons, J. E. Mazur, J. A. Nevin, & H. Rachlin (Eds.), *Quantitative analyses of behavior: Vol. V. The effect of delay and of intervening events on reinforcement value* (pp. 55–73). Hillsdale, NJ: Erlbaum.
- McKerchar, T. L., Green, L., Myerson, J., Pickford, T. S., Hill, J. C., & Stout, S. C. (2009). A comparison of four models of delay discounting in humans. *Behavioural Processes*, 81, 256–259.
- Myerson, J., & Green, L. (1995). Discounting of delayed rewards: Models of individual choice. *Journal of the Experimental Analysis of Behavior*, 64, 263–276.
- Myerson, J., Green, L., & Warusawitharana, M. (2001). Area under the curve as a measure of discounting. *Journal of the Experimental Analysis of Behavior*, 76, 235–243.
- Neef, N. A., Bicard, D. F., & Endo, S. (2001). Assessment of impulsivity and the development of self-control in students with attention deficit hyperactivity disorder. *Journal of Applied Behavior Analysis*, 34, 397–408.
- Ohmura, Y., Takahashi, T., Kitamura, N., & Wehr, P. (2006). Three-month stability of delay and probability discounting measures. *Experimental and Clinical Psychopharmacology*, 14, 318–328.
- Rachlin, H. (1989). *Judgment, decision, and choice: A cognitive/behavioral synthesis*. New York: Freeman.
- Rachlin, H., & Green, L. (1972). Commitment, choice and self-control. *Journal of the Experimental Analysis of Behavior*, 17, 15–22.

- Rachlin, H., Raineri, A., & Cross, D. (1991). Subjective probability and delay. *Journal of the Experimental Analysis of Behavior*, 55, 233–244.
- Reed, D. D., & Martens, B. K. (2008). Sensitivity and bias under conditions of equal and unequal academic task difficulty. *Journal of Applied Behavior Analysis*, 41, 39–52.
- Richards, J. B., Zhang, L., Mitchell, S. H., & de Wit, H. (1999). Delay or probability discounting in a model of impulsive behavior: Effect of alcohol. *Journal of the Experimental Analysis of Behavior*, 71, 121–143.
- Simpson, C. A., & Vuchinich, R. E. (2000). Reliability of a measure of temporal discounting. *The Psychological Record*, 50, 3–16.
- Vollmer, T. R., Borrero, J. C., Lalli, J. S., & Daniel, D. (1999). Evaluating self-control and impulsivity in children with severe behavior disorders. *Journal of Applied Behavior Analysis*, 32, 451–466.
- Wenar, C., & Kerig, P. (2006). *Developmental psychopathology: From infancy through adolescence* (5th ed.). Boston: McGraw Hill.

Received March 6, 2008

Final acceptance March 19, 2010

Action Editor, William Ahearn